

Масштабируемость CFD–приложений на гибридных вычислительных системах¹

А.А. Давыдов

Аннотация. Одной из важных характеристик вычислительных приложений является их масштабируемость, т.е. возможность ускоряться при увеличении объема задействованных вычислительных ресурсов. Эта характеристика относится не столько к самому приложению, сколько к совокупности «приложение – вычислительная система». С появлением мощных вычислительных систем гибридной архитектуры с большим числом графических процессоров (ГПУ) встает вопрос о возможности их эффективного использования для решения больших задач.

Ключевые слова: вычислительная газовая динамика, многопроцессорные вычислительные системы, графические процессоры, квазигазодинамических уравнения.

Программный комплекс «Экспресс-3D»

В работе исследуется параллельная эффективность программного комплекса «Экспресс-3D» [4] при решении уравнений Эйлера и системы квазигазодинамических уравнений [1] на вычислительных системах K-100 [6] и СК «Ломоносов» [7] с использованием графических процессоров NVIDIA Fermi (Tesla C20XX).

Программный комплекс «Экспресс-3D» [2, 3] предназначен для решения задач газовой динамики на многоблочных индексных сетках. На топологию сетки наложены следующие ограничения: блоки стыкуются вершина в вершину; сетка на границе блоков стыкуется узел в узел. На одном вычислительном устройстве (на ГПУ или ядре ЦПУ) могут обрабатываться как один, так и несколько блоков. В комплексе реализованы несколько математических моделей (системы уравнений Эйлера, Навье-Стокса, модель квазигазодинамических уравнений). Для всех математических моделей используются явные методы конечного объема. Расчет может проводиться как на графических процессорах,

так и на ядрах универсального процессора. В случае использования графических процессоров ядро универсального процессора, управляющее графическим процессором, выполняет только коммуникационные функции и непосредственно вычислений не проводит. Для обработки блока на ГПУ применяется технология CUDA. Блок обрабатывается как единое целое. Если на одно вычислительное устройство приходится несколько блоков, то они обрабатываются последовательно.

В независимости от того, где обрабатываются два соседних блока: на одном вычислительном устройстве или на разных, обмен данными между ними производится через промежуточные массивы. Это необходимо для учета взаимной ориентации соседствующих граней. Следовательно, подготовка данных для обменов (заполнение вспомогательных массивов и их последующая распаковка в рабочие массивы) не является накладными расходами в многопроцессорном варианте расчета, а составляет часть алгоритма решения задачи на многоблочной сетке. Таким образом, накладными расходами при параллельной реализации будут только пересылки

¹ Работа выполнена при поддержке РФФИ (проекты 11-08-00269-а, 12-01-00556-а)

данных между узлами вычислительной системы (сюда входят и пересылки между графическим и управляющим процессорами в случае обмена между ГПУ, расположенными на различных узлах вычислительной системы) и задержки на синхронизацию.

Обмен данными между узлами в ПК «Экспресс-3D» осуществляется посредством коммуникационной среды *shmem* с прямой одностронней передачей данных. Данные передаются асинхронно непосредственно после обработки каждого блока. После обработки всех блоков делается барьерная синхронизация.

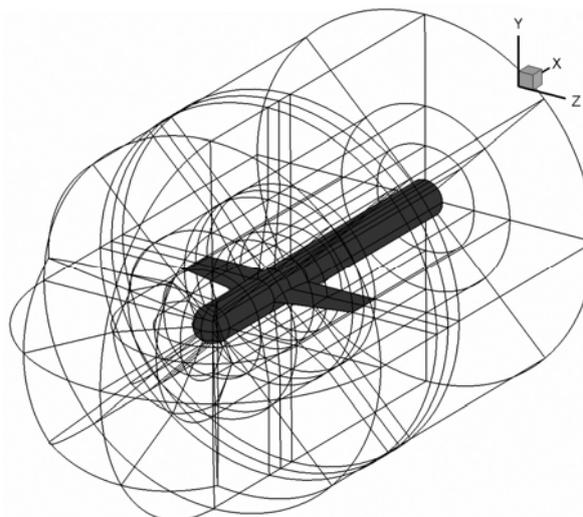
Масштабируемость приложений

Под масштабируемостью приложений, как правило, понимают одну из двух различных характеристик, каждая из которых по-своему важна.

Масштабируемость по числу процессоров при фиксированном объеме исходных данных — способность приложения ускоряться при увеличении числа процессоров и фиксированном объеме данных — $Eff = \frac{T(1)}{T(N) \cdot N} \%$, $T(1)$,

$T(N)$ — времена выполнения на одном и N процессорах соответственно. Как правило, размер конкретной прикладной задачи заранее известен и фиксирован. И в данном случае встает вопрос количества ресурсов, которые могут быть эффективно использованы для решения данной задачи. Эта характеристика по сути своей определяет характерный минимальный объем данных, который целесообразно обрабатывать на одном вычислительном устройстве или, как его еще называют, зерно параллелизма. Следует отметить, что для многоблочных сеток этот размер определяется структурой сетки и возможностью качественной балансировки загрузки (предполагается, что структура сетки остается постоянной и блоки на части не делятся).

Масштабируемость по числу процессоров с пропорциональным наращиванием объема исходных данных — способность приложения не замедляться при увеличении объема задачи и пропорциональном увеличении числа процессоров — $Eff = \frac{T(V_N)}{T(V_1)} \%$, $T(V_1)$, $T(V_N)$ — вре-



Многоблочная сетка около ЛА

мена обработки соответствующих объемов данных на одном и N процессорах. Эта характеристика показывает, какое максимальное число вычислительных устройств может быть эффективно задействовано и общий объем обрабатываемых при этом данных.

Здесь везде процессор понимается в общем смысле вычислительного устройства.

Методика эксперимента

Масштабируемость при фиксированном объеме исследовалась на задаче об обтекании крылатого тела сверхзвуковым потоком газа в рамках уравнений Эйлера. Сетка состояла из 132 блоков различной размерности. Общее число ячеек около 2 млн. Расчетная область и структура блоков изображена на рисунке.

Размер минимального блока — 4050 ячеек, максимального — 197579 ячеек. Задача разбивалась на различное число графических процессоров: 1, 3, 9, 27, соответственно. Кратность числа процессоров трем обусловлена составом ГСК К-100 — каждый узел содержит 3 ГПУ. Времена расчетов на различном числе графических процессоров приведены в Табл. 1.

Увеличение числа ГПУ свыше 27 для данной задачи нецелесообразно в силу плохой балансировки загрузки.

Масштабируемость с наращиванием объема данных исследовалась на хорошо известной задаче о течении в прямоугольной каверне с квад-

Табл. 1. Параллельная эффективность при постоянном размере задачи

Количество ГПУ (узлов на K-100)	1 (1)	3 (1)	9 (3)	27 (9)
Время	27,90	9,88	3,58	1,35
Ускорение	1,00	2,82	7,79	20,67
Эффективность	100,00 %	94,13 %	86,59 %	76,54 %

Табл. 2. Параллельная эффективность при пропорциональном увеличении размере задачи

Число GPU	1	2	4	8	16	32	64	128	256	512	1024
Время, с.	87,12	87,82	88,8	89,30	90,23	91,00	91,50	91,57	91,97	92,46	92,74
Эффективность %	100,0	99,2	98,1	97,6	96,6	95,7	95,2	95,1	94,7	94,2	93,9

Табл. 3. Времена совместного расчета на ГПУ + ЦПУ

	1 узел K-100			2 узла K-100		
	11 CPU	3 GPU	3GPU+8CPU	22 CPU	6GPU	6GPU+16CPU
Время	184.2	65.98	47.72	102	38.8	28.15
Ускорение относительно ЦПУ	1	2.79	3.86	1	2.62	3.62
Ускорение относительно ГПУ		1	1.38		1	1.37

ратной крышкой [4]. На каждом графическом процессоре обрабатывался один блок размерностью 150x150x150. Таким образом, суммарный размер сетки, обрабатываемый на 1024 ускорителях, составил 3.5 миллиарда ячеек. В Табл. 2 приведены времена расчета данной задачи на различном числе ГПУ на СК «Ломоносов».

За характерный объем задачи, относительно которого измеряется эффективность, был выбран объем, максимально приближенный к памяти одного вычислительного устройства (графического процессора).

Совместное использование ЦПУ и ГПУ

Современная элементная база, из которой собираются гибридные вычислительные системы, такова, что на один графический ускоритель приходится несколько ядер универсального процессора. Для управления каждым графическим процессором, чаще всего, отводится одно ядро универсального процессора, а остальные, обычно, не задействованы.

Комплекс «Экспресс-3D» позволяет использовать для расчета не только графические процессоры, но и ядра универсального процессора. Так, если узел вычислительной системы состоит

из 8 ядер универсального процессора и 2 графических ускорителей (например, СК «Ломоносов», МГУ), то работу следует разбивать на 8 частей — на 6 универсальных ядер и на 2 пары CPU+GPU. Конечно, сложность балансировки загрузки при этом существенно возрастает.

В работе [3] предложен эффективный алгоритм, позволяющий сбалансировать параллельную работу ГПУ и ЦПУ. Основная идея алгоритма состоит в том, чтобы мелкие блоки, которые не очень хорошо ускоряются на ГПУ (имеется ввиду отношение времен обработки блока на графическом процессоре и ядре универсального процессора), но число которых, как правило, велико, считать на ЦПУ. Применение такого подхода позволяет на 30–40% повысить суммарную эффективность узла вычислительной системы (по сравнению с вычислениями только на ускорителях). В Табл. 3 приведены времена расчетов упомянутой выше задачи об обтекании крылатого тела на различном числе ядер универсального процессора и графических процессорах.

Следует, однако, отметить, что такой подход будет эффективен только на сильно неоднородных по размерам блоках сетках и при делении на не очень большое число узлов.

Заключение

Показана высокая параллельная эффективность комплекса «Экспресс-3D» на вычислительных системах К-100 и СК «Ломоносов» как в смысле ядра параллелизма, так и в смысле максимального объема обрабатываемых данных.

Полученные результаты дают возможность прогнозировать достаточно высокую параллельную эффективность при решении задач газовой динамики на системах с достаточно большим числом (до нескольких десятков тысяч) графических процессоров, что соответствует системам с пиковой производительностью 1–10 Петафлопс.

Литература

1. Елизарова Т.Г. Квазигазодинамические уравнения и методы расчета вязких течений. М.: Научный мир, 2007.
2. А.А. Давыдов. Параллельный программный комплекс «Express-3D» для решения задач газовой динамики в областях сложной формы на гибридных вычислительных системах с графическими процессорами NVIDIA. Материалы XI всероссийской конференции «Высокопроизводительные параллельные вычисления на кластерных системах». Нижний Новгород, 2 ноября 2011 г., СС 100-101.
3. А.А. Давыдов. Численное моделирование задач газовой динамики на гибридных вычислительных системах. Диссертация на соискание ученой степени кандидата физико-математических наук. Москва 2012.
4. Zoran Zunic, Matjaz Hribersek, Leopold Skerget and Jure Ravnik. 3D lid driven cavity flow by mixed boundary And Finite Element Method. European Conference on Computational Fluid Dynamics, ECCOMAS CFD 2006.
5. Boris N. Chetverushkin, Alexander A. Davydov, Evgeny V. Shilnikov Numerical Simulation of the Continuous Media Problems on Hybrid Computer Systems. PARENG 2011, Ajaccio, Corsica, France, 12-15 April 2011
6. <http://www.kiam.ru/MVS/resources/k100.html>
7. <http://parallel.ru/cluster/lomonosov.html>

Давыдов Александр Александрович. Научный сотрудник Института прикладной математики им. М. В. Келдыша РАН. Окончил МГУ им. М. В. Ломоносова в 2002 году. Кандидат физико-математических наук. Область научных интересов: параллельные вычисления, вычислительные методы, газовая динамика, турбулентные течения, слабосжимаемые течения. E-mail: alexander.a.davydov@gmail.com